Fig 15 an antibody, wherein said antibody is specific for TADG-15 protein.

(Please amend claim 24 as follows:)

24. (thrice amended) An antibody, wherein said antibody is Fig 16 specific for Tumor Antigen Derived Gene-15 (TADG-15) protein.

REMARKS

Amendment

Attached hereto is a marked-up version of the changes made to the specification and claims by the current amendment. The attached page is captioned "Version with markings to show changes made."

Specification

Applicants submit that the amendments to the specification, including to the Brief Description of the Drawings, only amend the figure identifiers to correspond to the identifiers on the drawings submitted with the application; i.e., Figures 1, 2, 11, and 12 are amended to Figures 1A/1B, 2A/2B/2C/2D, 11A/11B, and 12A/12B/12C/12D/12E, and contain no new matter.

Priority

Claims 22 and 24 have been amended to recite an antibody as originally submitted; therefore, the priority of the instant application is the filing date of the parent application filed February 20, 1998.

Drawings

Corrected drawings are submitted herein.

The 35 U.S.C. §112 Rejection

Claims 22 and 24 were rejected under 35 U.S.C. §112, first paragraph, as containing new matter. The rejection is respectfully traversed.

Claims 22 and 24 have been amended to recite an antibody as originally submitted. Accordingly, Applicants respectfully request that the rejection of claims 22 and 24 under 35 U.S.C. §112, first paragraph, be withdrawn.

The Examiner stated that the 103(a) rejection set forth in the office action mailed June 29, 2000 would be reinstated if the claims are amended as originally presented. In the office action mailed June 29, 2000, claims 22-24 were rejected under 35 U.S.C. §103(a) as being

unpatentable over Accession #W22987 in view of Lerner. The rejection is respectfully traversed.

Accession #W22987 disclosed a polypeptide that has a sequence identical to part of the TADG-15 protein disclosed in the instant application. The Examiner argued that it would have been obvious to use part or all of the polypeptide of Accession #W22987 to produce antibodies reactive with TADG-15 protein. In the office action mailed July 13, 2001, the Examiner stated that an antibody raised against Accession #W22987 would fit the criterion listed in Applicants' claims. Applicants respectfully disagree.

The claims of the instant application are drawn to an antibody specific for TADG-15 protein. Antibodies raised against the polypeptide of Accession #W22987 are not specific for TADG-15 protein. These antibodies would bind to both the polypeptide of Accession #W22987 and TADG-15 protein, i.e. these antibodies are cross-reactive between the polypeptide of Accession #W22987 and TADG-15 protein rather than specific for TADG-15 protein.

In the office action mailed July 13, 2001, the Examiner stated that the term "specific" is not the same as exclusive. Applicants respectfully disagree. Applicants submit that one of ordinary skill in the art would understand and interpret an antibody specific for TADG-15

protein as an antibody that only binds to TADG-15 protein and does not shows cross-reactivity to other proteins. One of ordinary skill in the art would not take antibodies that cross-react between the polypeptide of Accession #W22987 and TADG-15 protein as antibodies that are specific for the TADG-15 protein. Hence, in contrast to the Examiner's assertion, the term "specific" is the same as exclusive as readily understood by one of ordinary skill in the art.

In view of the above remarks, Accession #W22987 and Lerner do not teach or suggest an antibody specific for TADG-15 protein. The invention as a whole is not *prima facie* obvious to one of ordinary skill in the art at the time the invention was made. Accordingly, Applicants respectfully request that the rejection of claims 22-24 under 35 U.S.C. §103(a) be withdrawn.

The 35 U.S.C. §102 Rejection

Claim 24 was rejected under 35 U.S.C. §102(a) as being anticipated by **Lin** et al. The rejection is respectfully traversed.

Claim 24 has been amended to recite an antibody as originally submitted, thereby claiming a priority date of February 20, 1998 which is before the publication date of **Lin** et al. Accordingly,

Applicants respectfully request that the rejection of claim 24 under 35 U.S.C. §102(a) be withdrawn.

The 35 U.S.C. §103(a) Rejection

Claims 22-24 were rejected under 35 U.S.C. §103(a) as being unpatentable over **Lin** et al. However, as stated above, claims 22-24 have a priority date before the publication date of **Lin** et al. Accordingly, Applicants respectfully request that the rejection of claims 22-24 under 35 U.S.C. §103(a) be withdrawn.

This is intended to be a complete response to the Office Action mailed April 3, 2002. If any issues remain outstanding, the Examiner is respectfully requested to telephone the undersigned attorney of record for immediate resolution.

Respectfully submitted,

Date: July 26, 2002

ADLER & ASSOCIATES
8011 Candle Lane
Houston, Texas 77071
(713) 270-5391
badler1@houston.rr.com

Benjamin Aaron Adler, Ph.D., J.D.
Registration No. 35,423
Counsel for Applicant

VERSION WITH MARKINGS TO SHOW CHANGES MADE

IN THE SPECIFICATION:

Paragraph beginning on line 19 of page 9 has been amended as follows:

Figures 1A and 1B show ~~**Figure 1 shows**~~ a comparison of the serine protease catalytic domain of TADG-15 with Hepsin (Heps. SEQ ID No. 3), SCCE (SEQ ID No. 4), Trypsin (Try, SEQ ID No. 5), Chymotrypsin (Chymb, SEQ ID No. 6), Factor 7 (Fac7, SEQ ID No. 7) and Tissue plasminogen activator (Tpa, SEQ ID No. 8). The asterisks indicate conserved amino acids of catalytic triad.

Paragraph beginning on line 10 of page 4 has been amended as follows:

Figures 2A-2D show ~~**Figure 2 shows**~~ the nucleotide sequence of the TADG-15 cDNA and the derived amino acid sequence of the TADG-15 protein. The putative start codon is located at nucleotides 23-25. The potential transmembrane sequence is underlined. Possible N-linked glycosylation sites are indicated by a broken line. The asterisks indicate conserved cysteine residues of CUB domain. The SDE-motifs of the LDL receptor ligand binding repeat-like domain are boxed. The arrow shows the arginine-valine bond cleaved upon activation. The

conserved amino acids of the catalytic triad; histidine, aspartic acid and serine residues are circled.

Paragraph beginning on line 19 of page 12 has been amended as follows:

Figures 11A and **11B** show ~~**Figure 11** shows~~ an alignment of the human TADG15 protein sequence with that of mouse epithin which demonstrates that the proteins re 84% similar and 81% identical over 843 amino acids. Residues that are identical between the two proteins are indicated by a "-", while the "*" symbol represents the TADG15 translation termination. The most significant difference between these two proteins lies in the carboxy-termini, which for epithin, includes 47 amino acids that are not present in TADG15.

Paragraph beginning on line 6 of page 13 has been amended as follows:

Figures 12A-12E show ~~**Figure 12** shows~~ a nucleotide sequence comparison between TADG-15 and human SNC-19 (GeneBank Accession No. #U20428).

Paragraph beginning on line 19 of page 33 has been amended as follows:

The invention includes a substantially pure DNA encoding a TADG-15 protein, a DNA strand which will hybridize at high stringency

to a probe containing a sequence of at least 15 consecutive nucleotides of (SEQ ID No.1). The protein encoded by the DNA of this invention may share at least 80% sequence identity (preferably 85%, more preferably 90%, and most preferably 95%) with the amino acids listed in Figures 3 and 4 (SEQ ID No. 2). More preferably, the DNA includes the coding sequence of the nucleotides of Figures 2A-2D ~~2~~ (SEQ ID No. 1), or a degenerate variant of such a sequence. This invention also includes a substantially pure DNA containing a sequence of at least 15 consecutive nucleotides (preferably 20, more preferably 30, even more preferably 50, and most preferably all) of the region from nucleotides 1 to 3147 of the nucleotides shown in Figures 2A-2D ~~2~~ (SEQ ID No. 1).

Paragraph beginnng on line 12 of page 34 has been amended as follows:

By "substantially pure DNA" is meant DNA that is not part of a milieu in which the DNA naturally occurs, by virtue of separation (partial or total purification) of some or all of the molecules of that milieu, or by virtue of alteration of sequences that flank the claimed DNA. The term therefore includes, for example, a recombinant DNA which is incorporated into a vector, into an autonomously replicating plasmid or virus, or into the genomic DNA of a prokaryote or eukaryote; or which exists as a separate molecule (*e.g.*, a cDNA or a genomic or cDNA

fragment produced by polymerase chain reaction (PCR) or restriction endonuclease digestion) independent of other sequences. It also includes a recombinant DNA which is part of a hybrid gene encoding additional polypeptide sequence, *e.g.*, a fusion protein. Also included is a recombinant DNA which includes a portion of the nucleotides listed in Figures 2A-2D ~~2~~ (SEQ ID No. 1) and which encodes an alternative splice variant of TADG-15.

Paragraph beginning on line 18 of page 37 has been amended as follows:

The probe to which the DNA of the invention hybridizes preferably consists of a sequence of at least 20 consecutive nucleotides, more preferably 40 nucleotides, even more preferably 50 nucleotides, and most preferably 100 nucleotides or more (up to 100%) of the coding sequence of the nucleotides listed in Figures 2A-2D ~~2~~ (SEQ ID No. 1) or the complement thereof. Such a probe is useful for detecting expression of TADG-15 in a cell by a method including the steps of (a) contacting mRNA obtained from the cell with a labeled TADG-15 hybridization probe; and (b) detecting hybridization of the probe with the mRNA.

Paragraph beginning on line 15 of page 38 has been amended as follows:

The DNA may have at least about 70% sequence identity to the coding sequence of the nucleotides listed in Figures 2A-2D ~~2~~ (SEQ ID No. 1), preferably at least 75% (*e.g.*, at least 80%); and most preferably at least 90%. The identity between two sequences is a direct function of the number of matching or identical positions. When a position in both of the two sequences is occupied by the same monomeric subunit, *e.g.*, if a given position is occupied by an adenine in each of two DNA molecules, then they are identical at that position. For example, if 7 positions in a sequence 10 nucleotides in length are identical to the corresponding positions in a second 10-nucleotide sequence, then the two sequences have 70% sequence identity. The length of comparison sequences will generally be at least 50 nucleotides, preferably at least 60 nucleotides, more preferably at least 75 nucleotides, and most preferably 100 nucleotides. Sequence identity is typically measured using sequence analysis software (*e.g.*, Sequence Analysis Software Package of the Genetics Computer Group (GCG), University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, WI 53705).

Paragraph beginning on line 4 of page 45 has been amended as follows:

Likewise, a standard Northern blot assay can be used to ascertain the relative amounts of TADG-15 mRNA in a cell or tissue

obtained from a patient suspected of having cancer, in accordance with conventional Northern hybridization techniques known to those of ordinary skill in the art. This Northern assay uses a hybridization probe, *e.g.*, radiolabelled TADG-15 cDNA, either containing the full-length, single stranded DNA having a sequence complementary to SEQ ID No. 1 (Figures 2A-2D ~~2~~), or a fragment of that DNA sequence at least 20 (preferably at least 30, more preferably at least 50, and most preferably at least 100 consecutive nucleotides in length). The DNA hybridization probe can be labeled by any of the many different methods known to those skilled in this art.

Paragraph beginning on line 14 of page 59 has been amended as follows:

A computerized search of GenEMBL databases using the FASTA program (Wisconsin Package Version 9.1, GCG, Madison, Wisconsin) for amino acid sequences homologous to the TADG-15 protease domain revealed that homologies with other known human proteases never exceeds 55%. Figures 1A and 1B show ~~1 shows~~ the alignment of the protease domain of TADG-15 compared with other human serine proteases. Using the BESTFIT program available through GCG, the similarities between TADG-15 and trypsin, chymotrypsin, and tissue-type plasminogen activator are 51%, 46% and 52%, respectively.

Paragraph beginning on line 3 of page 60 has been amended as follows:

From the sequence derived from the TADG-15 catalytic domain, specific primers were synthesized to amplify a TADG-15-specific probe for library screening. After screening an ovarian carcinoma library, one 1785 bp clone was obtained which included the 3' end of the TADG-15 transcript. Upon further screening using the 5' end of the newly detected clone, two additional clones were identified which provided another 1362 bp of the cDNA, including the 5' end of the TADG-15 transcript. The total length of the sequenced cDNA was approximately 3.15 kb. The total nucleotide sequence obtained includes a Kozak's consensus sequence preceding a single open reading frame encoding a predicted protein of 855 amino acids (Figures 2A-2D ~~2~~).

Paragraph beginning on line 15 of page 60 has been amended as follows:

The deduced open reading frame encoded by the TADG-15 nucleotide sequence (Figures 2A-2D ~~2~~, 3 and 4) contains several distinct domains as follows: an amino terminal cytoplasmic tail (amino acids (aa) # 1-54), a potential transmembrane domain (aa #55-77), an extracellular membrane domain (aa #78-213), two complement subcomponents Clr/Cls, Uegf, and bone morphogenetic protein 1 (CUB)

repeats (aa #214-447), four ligand binding repeats of the low density lipoprotein (LDL) receptor-like domain (aa #453-602) and a serine protease domain (aa #615-855). The TADG-15 protein also contains two potential N-linked glycosylation sites (aa #109 and 302) and a potential proteolytic cleavage site upstream from the protease domain (aa #614) which could release and/or activate the protease at the carboxy end of this protein. In addition, TADG-15 containsan RGD motif (aa #249-251) which is commonly found in proteins involved in cell-cell adhesion.

Paragraph beginning on line 4 of page 68 has been amended as follows:

Recently, a mouse protein named epithin (GenBank Accession No. AF042822) has been described.[14] Epithin is a 902 amino acid protein which contains a similar structure to TADG-15 in that it has a cytoplasmic domain, transmembrane domain, two CUB domains, four LDLR-like domains and a carboxy terminal serine protease domain. TADG-15 and epithin are 84% similar over 843 amino acids, suggesting that the proteins may be orthologous (Figures 11A and 11B ~~11~~). The precise role of epithin remains to be elucidated.

Paragraph beginning on line 12 of page 68 has been amended as follows:

A search of GeneBank for similar previously identified sequences yielded one such sequence with relatively high homology to a portion of the TADG-15 gene. The similarity between the portion of TADG-15 from nucleotide #182 to 3139 and SNC-19 GeneBank Accession No. #U20428) is approximately 97% (Figures 12A-12E ~~12~~). There are however significant differences between SNC-19 and TADG-15. For example, TADG-15 has an open reading frame of 855 amino acids whereas the longest open reading frame of SNC-19 is 173 amino acids. Additionally, SNC-19 does not include a proper start site for the initiation of translation, nor does it include the amino terminal portion of the protein encoded by TADG-15. Moreover, SNC-19 does not include an open reading frame for a functional serine protease because the His, Asp and Ser residues of the catalytic triad that are necessary for function are encoded in different reading frames.

IN THE CLAIMS:

Claim 22 has been amended as follows:

22. (thrice amended) A kit for detecting Tumor Antigen Derived Gene-15 (TADG-15) protein, comprising:

an antibody, wherein said antibody is specific for TADG-15 protein, ~~wherein said antibody is directed against all or part of amino acids 1 to 614 of SEQ ID No. 2~~.

Claim 24 has been amended as follows:

24. (thrice amended) An antibody, wherein said antibody is specific for ~~amino acids 1 to 615 of~~ Tumor Antigen Derived Gene-15 (TADG-15) protein, ~~wherein said antibody is directed against all or part of amino acids 1 to 614 of SEQ ID No. 2~~.

```
1   TCAAGAGCGGCCTCGGGGTACCATGGGGAGCGATCGGGCCCGCAAGGGCGGAGGGGGCCCGAAGGACTTCGGCGC
                      M  G  S  D  R  A  R  K  G  G  G  G  P  K  D  F  G  A  18
76  GGGACTCAAGTACAACTCCCGGCACGAGAAAGTGAATGGCTTGGAGGAAGGCGTGGAGTTCCTGCCAGTCAACAA
     G  L  K  Y  N  S  R  H  E  K  V  N  G  L  E  E  G  V  E  F  L  P  V  N  N  43
151 CGTCAAGAAGGTGGAAAAGCATGGCCCGGGGCGCTGGGTGGTGCTGGCAGCCGTGCTGATCGGCCTCCTCTTGGT
     V  K  K  V  E  K  H  G  P  G  R  W  V  V  L  A  A  V  L  I  G  L  L  V  68
226 CTTGCTGGGGATCGGCTTCCTGGTGTGGCATTTGCAGTACCGGGACGTGCGTGTCCAGAAGGTCTTCAATGGCTA
     L  L  G  I  G  F  L  V  W  H  L  Q  Y  R  D  V  R  V  Q  K  V  F  N  G  Y  93
301 CATGAGGATCACAAATGAGAATTTTGTGGATGCCTACGAGAACTCCAACTCCACTGAGTTTGTAAGCCTGGCCAG
     M  R  I  T  N  E  N  F  V  D  A  Y  E  N  S  N  S  T  E  F  V  S  L  A  S  118
376 CAAGGTGAAGGACGCGCTGAAGCTGCTGTACAGCGGAGTCCCATTCCTGGGCCCCTACCACAAGGAGTCGGCTGT
     K  V  K  D  A  L  K  L  L  Y  S  G  V  P  F  L  G  P  Y  H  K  E  S  A  V  143
451 GACGGCCTTCAGCGAGGGCAGCGTCATCGCCTACTACTGGTCTGAGTTCAGCATCCCGCAGCACCTGGTGGAGGA
     T  A  F  S  E  G  S  V  I  A  Y  Y  W  S  E  F  S  I  P  Q  H  L  V  E  E  168
526 GGCCGAGCGCGTCATGGCCGAGGAGCGCGTAGTCATGCTGCCCCCGCGGGCGCGCTCCCTGAAGTCCTTTGTGGT
     A  E  R  V  M  A  E  E  R  V  V  M  L  P  P  R  A  R  S  L  K  S  F  V  V  193
601 CACCTCAGTGGTGGCTTTCCCCACGGACTCCAAAACAGTACAGAGGACCCAGGACAACAGCTGCAGCTTTGGCCT
     T  S  V  V  A  F  P  T  D  S  K  T  V  Q  R  T  Q  D  N  S  C  S  F  G  L  218
676 GCACGCCCGCGGTGTGGAGCTGATGCGCTTCACCACGCCCGGCTTCCCTGACAGCCCCTACCCCGCTCATGCCCG
     H  A  R  G  V  E  L  M  R  G  T  T  P  G  R  P  D  S  P  Y  P  A  H  A  R  243
```

Fig. 2A

```
751 CTGCCAGTGGGCCCTGCGGGGGGACGCCGACTCAGTGCTGAGCCTCACCTTCCGCAGCTTTGACCTTGCGTCCTG
     C  Q  W  A  L  R  G  D  A  D  S  V  L  S  L  T  F  R  S  F  D  L  A  S  C  268
826 CGACGAGCGCGGCAGCGACCTGGTGACGGTGTACAACACCCTGAGCCCCATGGAGCCCCACGCCCTGGTGCAGTT
     D  E  R  G  S  D  L  V  T  V  Y  N  T  L  S  P  M  E  P  H  A  L  V  Q  L  293
901 GTGTGGCACCTACCCTCCCTCCTACAACCTGACCTTCCACTCCTCCCAGAACGTCCTGCTCATCACACTGATAAC
     C  G  T  Y  P  P  S  Y  N  L  T  F  H  S  S  Q  N  V  L  L  I  T  L  I  T  318
976 CAACACTGAGCGGCGGCATCCCGGCTTTGAGGCCACCTTCTTCCAGCTGCCTAGGATGAGCAGCTGTGGAGGCCG
     N  T  E  F  F  H  P  G  F  E  A  T  F  F  Q  L  P  R  M  S  S  C  G  G  R  343
1051 CTTACGTAAAGCCCAGGGGACATTCAACAGCCCCTACTACCCAGGCCACTACCCACCCAACATTGACTGCACATG
     L  R  K  A  Q  G  T  F  N  S  P  Y  Y  P  G  H  Y  P  P  N  I  D  C  T  W  368
1126 GAACATTGAGGTGCCCAACAACCAGCATGTGAAGGTGAGCTTCAAATTCTTCTACCTGCTGGAGCCCGGCGTGCC
     N  I  E  V  P  N  N  Q  H  V  K  V  S  F  K  F  F  Y  L  L  E  P  G  V  P  393
1201 TGCGGGCACCTGCCCCAAGGACTACGTGGAGATCAATGGGGAGAAATACTGCGGAGAGAGGTCCCAGTTCGTCGT
     A  G  T  C  P  K  D  Y  V  E  I  N  G  E  K  Y  C  G  E  R  S  Q  F  V  V  418
1276 CACCAGCAACAGCAACAAGATCACAGTTCGCTTCCACTCAGATCAGTCCTACACCGACACCGGCTTCTTAGCTGA
     T  S  N  S  N  K  I  T  V  R  F  H  S  D  Q  S  Y  T  D  T  G  F  L  A  E  443
1351 ATACCTCTCCTACGACTCCAGTGACCCATGCCCGGGGCAGTTCACGTGCCGCACGGGGCGGTGTATCCGGAAGGA
     Y  L  S  Y  D  S  S  D  P  C  P  G  Q  F  T  C  R  T  G  R  C  I  R  K  E  468
1426 GCTGCGCTGTGATGGCTGGGCCGACTGCACCGACCACAGCGATGAGCTCAACTGCAGTTGCGACGCCGGCCACCA
     L  R  C  D  G  W  A  D  C  T  D  H  S  D  E  L  N  C  S  C  D  A  G  H  Q  493
```

Fig. 2B

```
1501 GTTCACGTGCAAGAACAAGTTCTGCAAGCCCCTCTTCTGGGTCTGCGACAGTGTGAACGACTGCGGAGACAACAG
      F  T  C  K  N  K  F  C  K  P  L  F  W  V  C  D  S  V  N  D  C  G  D  N  S  518
1576 CGACGAGCAGGGGTGCAGTTGTCCGGCCCAGACCTTCAGGTGTTCCAATGGGAAGTGCCTCTCGAAAAGCCAGCA
      D  E  Q  G  C  S  C  P  A  Q  T  F  R  C  S  N  G  K  C  L  S  K  S  Q  Q  543
1651 GTGCAATGGGAAGGACGACTGTGGGGACGGGTCCGACGAGGCCTCCTGCCCCAAGGTGAACGTCGTCACTTGTAC
      C  N  G  K  D  D  C  G  D  G  S  D  E  A  S  C  P  K  V  N  V  V  T  C  T  568
1726 CAAACACACCTACCGCTGCCTCAATGGGCTCTGCTTGAGCAAGGGCAACCCTGAGTGTGACGGGAAGGAGGACTG
      K  H  T  Y  R  C  L  N  G  L  C  L  S  K  G  N  P  E  C  D  G  K  E  D  C  593
1801 TAGCGACGGCTCAGATGAGAAGGACTGCGACTGTGGGCTGCGGTCATTCACGAGACAGGCTCGTGTTGTTGGGGG
      S  D  G  S  D  E  K  D  C  D  C  G  L  R  S  F  T  R  Q  A  R  V  V  G  G  618
1876 CACGCATGCGGATGAGGGCGAGTGGCCCTGGCAGGTAAGCCTGCATGCTCTGGGCCAGGGCCACATCTGCGGTGC
      T  D  A  D  E  G  E  W  P  W  Q  V  S  L  H  A  L  G  Q  G  H  I  C  G  A  643
1951 TTCCCTCATCTCTCCCAACTGGCTGGTCTCTGCCGCACACTGCTACATCGATGACAGAGGATTCAGGTACTCAGA
      S  L  I  S  P  N  W  L  V  S  A  A  (H) C  Y  I  D  D  R  G  F  R  Y  S  D  668
2026 CCCCACGCAGTGGACGGCCTTCCTGGGCTTGCACGACCAGAGCCAGCGCAGCGCCCCTGGGGTGCAGGAGCGCAG
      P  T  Q  W  T  A  F  L  G  L  H  D  Q  S  Q  R  S  A  P  G  V  Q  E  R  R  693
2101 GCTCAAGCGCATCATCTCCCACCCCTTCTTCAATGACTTCACCTTCGACTATGACATCGCGCTGCTGGAGCTGGA
      L  K  R  I  I  S  H  P  F  F  N  D  F  T  F  D  Y (D) I  A  L  L  E  L  E  718
2176 GAAACCGGCAGAGTACAGCTCCATGGTGCGGCCCATCTGCCTGCCGGACGCCTCCCATGTCTTCCCTGCCGGCAA
      K  P  A  E  Y  S  S  M  V  R  P  I  C  L  P  D  A  S  H  V  F  P  A  G  K  743
```

Fig. 2C


```
2251 GGCCATCTGGGTCACGGGCTGGGGACACACCCAGTATGGAGGCACTGGCGCGCTGATCCTGCAAAAGGGTGAGAT
     A  I  W  V  T  G  W  G  H  T  Q  Y  G  G  T  G  A  L  I  L  Q  K  G  E  I  768
2326 CCGCGTCATCAACCAGACCACCTGCGAGAACCTCCTGCCGCAGCAGATCACGCCGCGCATGATGTGCGTGGGCTT
     R  V  I  N  Q  T  T  C  E  N  L  L  P  Q  Q  I  T  P  R  M  M  C  V  G  F  793
2401 CCTCAGCGGCGGCGTGGACTCCTGCCAGGGTGATTCCGGGGGACCCCTGTCCAGCGTGGAGGCGGATGGGCGGAT
     L  S  G  G  V  D  S  C  Q  G  D  (S) G  G  P  L  S  S  V  E  A  D  G  R  I  818
2476 CTTCCAGGCCGGTGTGGTGAGCTGGGGAGACGGCTGCGCTCAGAGGAACAAGCCAGGCGTGTACACAAGGCTCCC
     F  Q  A  F  C  C  S  W  G  D  G  C  A  Q  R  N  K  P  G  V  Y  T  R  L  P  843
2551 TCTGTTTCGGGACTGGATCAAAGAGAACACTGGGGTATAGGGGCCGGGGCCACCCAAATGTGTACACCTGCGGGG
     L  F  R  D  W  I  K  E  N  T  G  V                (SEQ ID NO: 2)              855
2626 CCACCCATCGTCCACCCCAGTGTGCACGCCTGCAGGCTGGAGACTGGACCGCTGACTGCACCAGCGCCCCCAGAA
2701 CATACACTGTGAACTCAATCTCCAGGGCTCCAAATCTGCCTAGAAAACCTCTCGCTTCCTCAGCCTCCAAAGTGG
2776 AGCTGGGAGGTAGAAGGGGAGGACACTGGTGGTTCTACTGACCCAACTGGGGGCAAAGGTTTGAAGACACAGCCT
2851 CCCCCGCCAGCCCCAAGCTGGGCCGAGGCGCGTTTGTGTATATCTGCCTCCCCTGTCTGTAAGGAGCAGCGGGAA
2926 CGGAGCTTCGGAGCCTCCTCAGTGAAGGTGGTGGGGCTGCCGGATCTGGGCTGTGGGGCCCTTGGGCCACGCTCT
3001 TGAGGAAGCCCAGGCTCGGAGGACCCTGGAAAACAGACGGGTCTGAGACTGAAATTGTTTTACCAGCTCCCAGGG
3076 TGGACTTCAGTGTGTGTATTTGTGTAAATGGGTAAAACAATTTATTTCTTTTTAAAAAAAAAAAAAAAAAA
                                                              (SEQ ID NO: 1)
```

⌊⌋ : KOZAK'S CONSENSUS SEQUENCE
▭ : TRANSMEMBRANE DOMAIN
○ : CONSERVED AMINO ACIDS OF CATALYTIC TRIAD H,D,S

Fig. 2D

```
                                                  *
  Heps  RIVGGRDTSL GRWPWQVSL. ....RYDG.A HLCGGSLLSG DWVLTAAHCF PE....RNRV
Tadg15  RVVGGTDADE GEWPWQVSL. ....HALGQG HICGASLISP NWLVSAAHCY IDDRGFRYSD
  Scce  KIIDGAPCAR GSHPWQVAL. ....LSGNQL H.CGGVLVNE RWVLTAAHC. .........K
   Try  KIVGGYNCEE NSVPYQVSL. ....NSGYHF ..CGGSLINE QWVVSAGHC. .........Y
 Chymb  RIVNGEDAVP GSWPWQVSL. ....QDKTGF HFCGGSLISE DQVVTAAHC. .........GV
  Fac7  RIVGGKVCPK GECPWQVLL. ....LVNG.A QLCGGTLINT IWVVSAAHCF DKIKNWRNLI
   Tpa  RIKGGLFADI ASHPWQAAIF AKHRRSPGER FLCGGILISS CWILSAAHCF QERFPPHHL.
                                                     *
  Heps  LSRWRVFAGA VAQASPHGLQ LGVQAVVYHG GYLPFRDPNS EENSNDIALV HLSS.PLPLT
Tadg15  PTWETAFLHL HDQSQRSAPG VQERRLKRII SHPFFNDFTF D...YDIALL ELEK.FAEYS
  Scce  MNEYTVHLGS DTLG..DR.R AQRIKASKSF RHPGYSTQT. ..HVNDLMLV KLNS.QARLS
   Try  KSRIQVRLGE HNIEVLEG.N EQFINAAKII RHPQYDRKT. ..LNNDIMLI KLSS.RAVIN
 Chymb  RTSDVVVAGE FDQGSDEE.N IQVLKIAKVF KNPKFSILT. ..VNNDITLL KLAT.PARFS
  Fac7  .....AVLGE HDLSEHDGDE QSRRVAQVII P....STYVP GTTNHDIALL RLHQ.PVVLT
   Tpa  ....TVILGR .TYRVVPGEE EQKFEVEKYI VHKEFDDDTY D...NDIALL QLKSDSSRCA

  Heps  EYIQPVCLPA ...AGQALVD GKICTVTGWG NTQYYGQQ.A GVLQEAAVPI ISNDVCNGAD
Tadg15  SMVRPICLPD ...ASHVFPA GKAIWVTGWG HTQYGGTG.A LILQKGEIRV INQTTCE..N
  Scce  SMVKKVRLPS ...RCE..PP GTTCTVSGWG TTTSPDVTFP SDLMCVDVKL ISPQDCTKV.
   Try  ARVSTISLPT ...APP..AT GTKCLISGWG NTASSGADYP DELQCLDAPV LSQAKCEAS.
 Chymb  QTVSAVCLPS ...ADDDFPA GTLCATTGWG KTKYNANKTP DKLQQAALPL LSNAECKKS.
  Fac7  DHVVPLCLPE RTFSERTLAF VRFSLVSGWG QLLDRGATAL ELMVLNVPRL NTQDCLQQSR
   Tpa  QESSVVRTVC LPPADLQLPD WTECELSGYG KHEALSPFYS ERLKEAHVRL YPSSRCTSQH
```

Fig. 1A

```
          *
Heps   FYGN..QIKP KMFCAGYPEG G......IDA CQGDSGGPFV CEDSISRTPR WRLCGIVSWG
Tadg15 LLPQ..QITP RMMCVGFLSG G......VDS CQGDSGGPL. ..SSVEADGR IFQAGVVSWG
Scce   .YKD..LLEN SMLCAGIPDS K......KNA CNGDSGGPLV C....R.... GTLQGLVSWG
Try    .YPG..KITS NMFCVGFLEG G......KDS CQGDSGGPVV C....M.... GQLQGVVSWG
Chymb  .WGR..RITD VMICAG..AS G......VSS CMGDSGGPLV C....QKDGA WTLVGIVSWG
Fac7   KVGDSPNITE YMFCAGYSDG S......KDS CKGDSGGP.. ..HATHYRGT WYLTGIVSWG
Tpa    LLNRT..VTD NMLCAGDTRS GGPQANLHDA CQGDSGGPLV CLN....DGR MTLVGIISWG

Heps   T.GCALAQKP GVYTKVSDFR EWIFQAIKTH SEASGXVTQL --   (SEQ ID NO: 3)
Tadg15 D.GCAQRNKP GVYTRLPLFR DWIKENTGV- ---------- --   (SEQ ID NO: 14)
Scce   TFPCGQPNDP GVYTQVCKFT KWINDTMKKH R--------- --   (SEQ ID NO: 4)
Try    D.GCAQKNKP GVYTKVYNYV KWIKNTIAAN S--------- --   (SEQ ID NO: 5)
Chymb  DSTCS.TSSP GVYARVTKLI PWVQKILAAN ---------- --   (SEQ ID NO: 6)
Fac7   Q.GCATVGHF GVYTRVSQYI EWLQKLMRSE PRPGVLLRAP FP   (SEQ ID NO: 7)
Tpa    .LGCGQKDVP GVYTKVTNYL DWIRDNMRP- ---------- --   (SEQ ID NO: 8)
```


Fig. 1B

```
1    MGSDRARKGG GGPKDFGAGL KYNSRHEKVN GLEEGVEFLP VNNVKKVEKH   1
51   GPGRWVVLAA VLIGLLLVLL GIGFLVWHLQ YRDVRVQKVF NGYMRITNEN   2
101  FVDAYENSNS TEFVSLASKV KDALKLLYSG VPFLGPYHKE SAVTAFSEGS
151  VIAYYWSEFS IPQHLVEEAE RVMAEERVVM LPPRARSLKS FVVTSVVAFP
201  TDSKTVQRTQ DNSCSFGLHA RGVELMRFTT PGFPDSPYPA HARCQWALRG
251  DADSVLSLTF RSFDLASCDE RGSDLVTVYN TLSPMEPHAL VQLCGTYPPS
301  YNLTFHSSQN VLLITLITNT ERRHPGFEAT FFQLPRMSSC GGRLRKAQGT   3
351  FNSPYYPGHY PPNIDCTWNI EVPNNQHVKV SFKFFYLLEP GVPAGTCPKD
401  YVEINGEKYC GERSQFVVTS NSNKITVRFH SDQSYTDTGF LAEYLSYDSS
451  DPCPGQFTCR TGRCIRKELR CDGWADCTDH SDELNCSCDA GHQFTCKNKF
501  CKPLFWVCDS VNDCGDNSDE QGCSCPAQTF RCSNGKCLSK SQQCNGKDDC   4
551  GDGSDEASCP KVNVVTCTKH TYRCLNGLCL SKGNPECDGK EDCSDGSDEK
601  DCDCGLRSFT RQARVVGGTD ADEGEWPWQV SLHALGQGHI CGASLISPNW
651  LVSAAHCYID DRGFRYSDPT QWTAFLGLHD QSQRSAPGVQ ERRLKRIISH
701  PFFNDFTFDY DIALLELEKP AEYSSMVRPI CLPDASHVFP AGKAIWVTGW   5
751  GHTQYGGTGA LILQKGEIRV INQTTCENLL PQQITPRMMC VGFLSGGVDS
801  CQGDSGGPLS SVEADGRIFQ AGVVSWGDGC AQRNKPGVYT RLPLFRDWIK
851  ENTGV (SEQ.ID NO: 2)
```

* : **Conserved cysteine residue**

NXT : **Possible N-linked glycosylation site**

SDE : **Conserved SDE motif**

▼ : **Potential cleavage site**

O : **Conserved amino acids of catalytic triad H, D, S**

1. **Cytoplasmic domain**
2. **Transmembrane domain**
3. **CUB repeat**
4. **Ligand-binding repeat (class A motif) of LDL receptor like domain**
5. **Serine protease**

Fig. 3

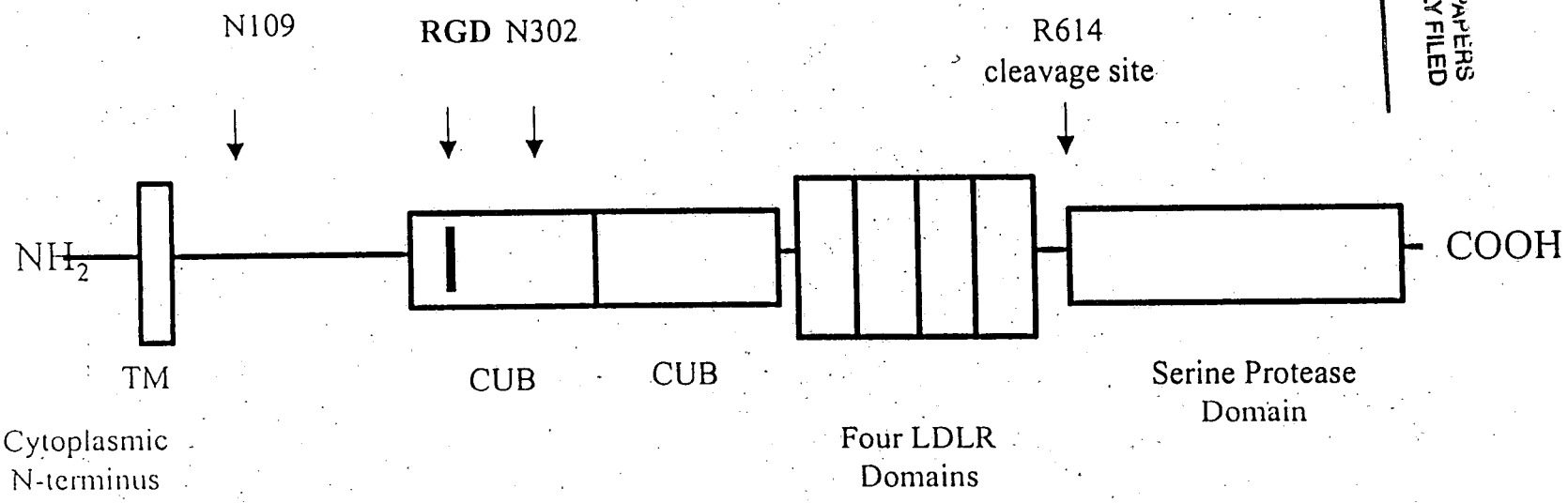

Fig. 4